Ornithokrites

Lukasz Tracewski
February 20, 2014

1 OVERVIEW

Ornithokrites is a transliteration of ancient Greek ὀρνιϊθοκρίτης, meaning interpreter of flight or cries of birds. With its rather ambitious name, the program itself is a tool meant for the automatic identification of kiwi calls from audio recordings. It is designed to cope with large variations of environmental conditions and low quality of input data. For each provided audio file, the program tries to find whether it contains any kiwi calls and, if so, whether they are male, female or both.

Complete source code can be found on project's web site: `https://github.com/tracek/Ornithokrites`.

2 HOW TO USE IT

Expected input are monaural (single-channel) audio files in Waveform Audio File Format (commonly known as WAV or WAVE). Following sections explain two ways of running the program: user-friendly 2.1 and user-hostile 2.2.

2.1 WEB INTERFACE

If the data is stored on Amazon Web Services S3 bucket, then by far easiest way of using the program is through a password-protected web site: `http://kiwi-finder.info`. The protection is needed since only one user at a time can run the program.

After providing the credentials user is directed to a simple web form that serves as an interface to the application.

- Bucket name: name of Amazon Web Services S3 bucket, e.g. *kiwicalldata.*

- Execute: connect to data store, download the recordings and run kiwi calls identification. It is a long-lasting operation. Closing the web page does not stop execution.

- Report: show results. Since they are generated live, user can click the button at any moment to get current state of affairs. Only text is printed, making it very fast.

- Show details: show detailed results. In this mode additional data is provided: spectrogram with identified fragments and option to play the original audio, allowing user to verify program's predictions.

- Clear: stop execution of the program and clear all intermediate results.

2.2 INTERACTIVE MODE

The program is written in Python, which means running it directly, either from command line or in interactive mode, requires installation of all dependent modules; complete list can be found on project's page. Due to the dependencies installing is not an easy task.

2.2.1 Batch mode - command line

`ornithiokrites.py -path [path to data]` - will run the identification on all WAVE files contained in `path` and its sub-folders.

2.2.2 Single-file mode - graphical user interface

If no command line arguments are provided then program will start in interactive mode. With open file dialog user can select a single file for analysis.

3 HOW IT WORKS

After the recordings are ready following steps take place:

1. **Apply high-pass filter**. This step will reduce strength of any signal below 1500 Hz. Experiments so far have shown that kiwi rarely show any vocalization below this value. It also helps to eliminate bird calls of no interest to us, e.g. long-tailed cuckoo.

2. **Find Regions of Interest** (ROIs), defined as any signal different than background noise. Since length of a single kiwi call is roughly constant, ROI length is fixed to one second. First onsets are found by calculating local energy of the input spectral frame and taking those above certain dynamically-assessed threshold. Then from the detected onset a delay of -0.2s is taken to compensate for possible discontinuities. End of ROI is defined as $+0.8$s after beginning of the onset, summing to 1s interval. The algorithm is made sensitive, since potential cost of not including kiwi candidate in set of ROIs is much higher then adding noise-only ROI.

3. **Reduce noise**. Since ROIs are identified, Noise-Only Regions (NORs) can be estimated as anything outside ROIs plus some margin. Based on NORs spectral subtraction is performed: knowing noise spectrum we can try to eliminate noise over whole sample.

4. **Calculate Audio Features**. Those features will serve as a kiwi audio signature, allowing to discriminate kiwi male from female - and the two from not a kiwi. For each ROI following features are calculated:

   - spectral flatness
   - perceptual spread
   - spectral roll-off
   - spectral decrease
   - spectral shape statistics
   - spectral slope
   - Linear Predictive Coding (LPC)
   - Line Spectral Pairs (LSP)

   AFs are calculated with `Yaafe` library. On its project page `http://yaafe.sourceforge.net/features.html` a complete description of above-mentioned features can be found.

5. **Perform kiwi identification**. At this stage Audio Features are extracted from the recording. Based on those, a Machine Learning algorithm, that is Support Vector Machines (SVM), will try to classify ROI as kiwi male, kiwi female and not a kiwi. At this moment additional rules are applied, employing our knowledge on repetitive character of kiwi calls. Only once sufficiently long set of calls is identified, the kiwi presence is marked.

6. **Report**. Algorithm output can be: female, male, male and female and no kiwi detected.

4 VALIDATION RESULTS

Program was tested using stratified 5-fold cross-validation. Based on provided training it has 98% accuracy in telling kiwi apart from non-kiwi and 92% accuracy.

	Kiwi Male	Kiwi Female	Male and Female	Not a kiwi
Kiwi Male	4	1	0	0
Kiwi Female	0	5	0	0
Male and Female	0	2	3	0
Not a kiwi	0	2	3	0

5 ACKNOWLEDGMENTS

that supplied me with invaluable help concerning kiwi identification. Without his help I would not be able to achieve such high accuracy.